



ENTERED

OIPF

## RAW SEQUENCE LISTING

DATE: 03/25/2002

PATENT APPLICATION: US/09/997,868

TIME: 15:23:32

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

## SEQUENCE LISTING

## W--&gt; 3 SEQUENCE LISTING

## 5 (1) GENERAL INFORMATION:

7 (i) APPLICANT: Gorman, Cornelia M.,

8 Groskreutz, Debyra J.

10 (ii) TITLE OF INVENTION: Prohormone Convertase Transformed Cells and  
11 Polypeptide Synthesis

13 (iii) NUMBER OF SEQUENCES: 57

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: Genentech, Inc.

17 (B) STREET: 1 DNA Way

18 (C) CITY: South San Francisco

19 (D) STATE: California

20 (E) COUNTRY: USA

21 (F) ZIP: 94080

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: WinPatin (Genentech)

29 (vi) CURRENT APPLICATION DATA:

C--&gt; 30 (A) APPLICATION NUMBER: US/09/997,868

C--&gt; 31 (B) FILING DATE: 12-Mar-2002

32 (C) CLASSIFICATION:

42 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 07/887265

36 (B) FILING DATE: 22-MAY-1992

39 (A) APPLICATION NUMBER: 07/803631

40 (B) FILING DATE: 06-DEC-1992

43 (A) APPLICATION NUMBER: PCT/US92/10621

44 (B) FILING DATE: 04-DEC-1992

46 (viii) ATTORNEY/AGENT INFORMATION:

47 (A) NAME: Love, Richard B.

48 (B) REGISTRATION NUMBER: 34,659

49 (C) REFERENCE/DOCKET NUMBER: P0748P3

51 (ix) TELECOMMUNICATION INFORMATION:

52 (A) TELEPHONE: 650/225-5530

53 (B) TELEFAX: 650/952-9881

54 (2) INFORMATION FOR SEQ ID NO: 1:

56 (i) SEQUENCE CHARACTERISTICS:

57 (A) LENGTH: 2355 base pairs

58 (B) TYPE: Nucleic Acid

59 (C) STRANDEDNESS: Single

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/997,868

DATE: 03/25/2002

TIME: 15:23:32

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

60 (D) TOPOLOGY: Linear  
62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
65 TCTAGATCTA GCTGGTGTGT CTCTGATCTT GCTTCTTTTC TCCCAGCCCT 50  
67 TCCTACTTGT GTGAGAACAA GGTTTTGAGC CATGGAGCAA AGAGGTTGGA 100  
69 CTCTGCAGTG TACTGCTTTC GCCTTCTTTT GCGTTTGGTG TGCACTAAGC 150  
71 AGTGTAAGAG CAAAGAGGCA GTTTGTAAAT GAATGGGCGG CGGAGATCCC 200  
73 CGGAGGGCAA GAAGCTGCCT CTGCCATCGC CGAAGAACTG GGGTATGACC 250  
75 TTTTGGGTCA GATTGGATCA CTTGAAAATC ACTATTTATT CAAACACAAA 300  
77 AGCCATCCTC GGAGGTCCCG AAGAAGCGCT CTTTATATCA CTAAGAGGTT 350  
79 ATCTGATGAT GATCGTGTGA CGTGGGCTGA ACAACAGTAT GAAAAAGAGA 400  
81 GAAGTAAACG TTCAGTTCAA AAAGACTCAG CATTGGATCT CTTCAATGAT 450  
83 CCAATGTGGA ATCAGCAGTG GTACTTGCAA GATACCAGAA TGAATGCAGC 500  
85 TCTGCCAAG CTGGACCTTC ATGTAATACC TGTTTGGGAA AAGGGTATTA 550  
87 CTGGCAAAGG AGTTGTTATT ACTGTACTGG ATGATGGCTT GGAGTGGAAT 600  
89 CACACAGACA TTTATGCCAA TTATGATCCA GAGGCTAGCT ATGATTTTAA 650  
91 CGATAATGAT CATGATCCAT TTCCCGGATA TGATCTCACA AATGAAAACA 700  
93 AACATGGAAC AAGATGTGCA GGTGAAATTG CCATGCAAGC AAATAATCAC 750  
95 AAGTGTGGGG TTGGAGTTGC ATATAATTCC AAAGTTGGAG GCATAAGAAT 800  
97 GCTGGATGGC ATTGTAATG ATGCCATTGA GGCTAGTTCA ATTGGATTCA 850  
99 ACCCTGGCCA TGTGGATATT TACAGTGCAA GCTGGGGCCC TAATGATGAT 900  
101 GGAATAAAGT TGGAGGGGCC TGGCAGACTA GCCCAGAAGG CATTGGAATA 950  
103 TGGTGTCAA CAGGGGAGAC AAGGGAAAGG CTCCATCTTT GTCTGGGCTT 1000  
105 CAGGGAATGG GGGTCGTCAG GGAGATAACT GTGACTGTGA TGGCTACACA 1050  
107 GACAGCATT ACACCATCTC TATCAGCAGT GCCTCCAGC AAGGCCTGTC 1100  
109 ACCTTGGTAT GCAGAGAAGT GTTCTTCCAC ATTGGCTACC TCCTACAGCA 1150  
111 GTGGTGATTA CACAGACCAG CGAATAACAA GCGCTGACCT GCACAATGAC 1200  
113 TGCACAGAGA CCCACACAGG CACCTCGGCT TCAGCACCCC TGGCTGCTGG 1250  
115 TATCTTTGCT CTGGCCTTGG AGGCAAACCC AAATCTTACC TGGAGAGATA 1300  
117 TGCAGCATCT GGTGTGTCTG ACCTCTGAGT ACGACCCATT GGCCAGTAAC 1350  
119 CCAGGTGGA AAAAGAATGG GGCAGGCTTG ATGGTGAACA GCCGATTGTTG 1400  
121 ATTTGGCTTG CTAATGCCA AAGCTCTGGT GGATTGCTG GATCCTCGGA 1450  
123 CCTGGAGAAA TGTGCCTGAG AAGAAAGAAT GTGTGTAAA AGACAATAAC 1500  
125 TTTGAGCCTA GAGCCCTGAA AGCTAATGGA GAAGTAATTG TTGAAATCCC 1550  
127 AACAAAGAGT TGTGAAGGAC AAGAAAATGC TATCAAGTCT CTGGAACATG 1600  
129 TGCAATTTGA AGCAACAATT GAATATTCTC GTAGAGGAGA CCTTCATGTC 1650  
131 AACTCACTT CTGCTGTTGG AACCAGCACT GTACTGTTGG CTGAAAGGGA 1700  
133 AAGAGATACA TCCCCAATG GCTTTAAGAA TTGGGACTTC ATGTCTGTTC 1750  
135 ATACATGGGG AGAGAATCCT GTAGGCACCT GGACATTGAA AATTACAGAC 1800  
137 ATGTCTGGAA GAATGCAAAA TGAAGGAAGG ATTGTGAAT GGAAGTTGAT 1850  
139 TTTGCATGGG ACATCTTCTC AACCAGAGCA CATGAAGCAG CCCCCTGTGT 1900  
141 ACACATCCTA CAATACAGTC CAGAATGACA GGAGAGGAGT GGAAGATG 1950  
143 GTGAATGTTG TGGAGAAGCG GCCACACAA AAGAGCCTGA ATGGCAATCT 2000  
145 CTGGTACCC AAAAATCCA GCAGCAGCAA TGTGGAGGT AGAAGGGATG 2050  
147 AGCAGGTACA AGGAATCCT TCAAAGGCCA TGCTGCGACT CCTACAAAGT 2100  
149 GCTTTTAGCA AGAATGCACT TTCAAACAA TCACCAAAGA AGTCTCCAAG 2150  
151 TGCAAAGCTC AGCATCCCTT ATGAAAGTTT CTATGAAGCC TTGGAAAAGC 2200  
153 TTAACAAGCC CTCCAAGCTT GAAGGCTCTG AAGACAGTCT GTACAGTGAC 2250  
155 TATGTTGATG TATTCTATA CACAAAACCT TATAAGCATA GAGATGACAG 2300  
157 GCTGCTGCAA GCTCTCATGG ACATCCTAAA TGAGGAGAAT TAAAATAAGG 2350

## RAW SEQUENCE LISTING

DATE: 03/25/2002

PATENT APPLICATION: US/09/997,868

TIME: 15:23:32

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

```

159  AGCTC 2355
161  (2) INFORMATION FOR SEQ ID NO: 2:
163      (i) SEQUENCE CHARACTERISTICS:
164          (A) LENGTH: 2012 base pairs
165          (B) TYPE: Nucleic Acid
166          (C) STRANDEDNESS: Single
167          (D) TOPOLOGY: Linear
169      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
172  TCTAGATGCA TCTTCCCTCT TCGTCCCCTG CTCCACCACC CTGCGCGCCT 50
174  CACAGCCCCG CTTTTCATCT CCAAAGAAGG ATGGAGGGCG GTTGTGGATC 100
176  CCAGTGGGAG GCGGCCGGGT TCCTCTTCTG TGTGATGGTT TTGCGTCTG 150
178  CCGAGAGACC CGTCTTCACG AATCATTTTC TTGTGGAGTT GCATAAAGAC 200
180  GGAGAGGAAG AGGCTCGCCA AGTTGCAGCA GAACACGGCT TTGGAGTCCG 250
182  AAAGTCCCCC TTTGCAGAAG GCCTGTATCA CTTTATCAC AATGGGCTTG 300
184  CAAAGGCCAA AAGAAGACGC AGCCTACACC ATAAGCGGCA GCTAGAGAGA 350
186  GACCCAGGA TAAAGATGGC GCTGCAACAA GAAGGATTTG ACCGTAAAAA 400
188  GAGAGGGTAC AGGGACATCA ATGAGATTGA CATCAACATG AATGATCCTC 450
190  TCTTTACAAA GCAATGGTAC CTGTTCAACA CTGGGCAAGC CGATGGAAC 500
192  CCTGGGCTAG ACTTGAACGT GGCCGAAGCC TGGGAGCTGG GATACACAGG 550
194  AAAAGGAGTG ACCATTGGAA TTATGGATGA TGGGAATTGAC TATCTCCACC 600
196  CAGACCTGGC CTACAACCTAC AACGCTGATG CAAGTTATGA CTTCAGCAGC 650
198  AATGACCCCT ACCCATACCC TCGATACACA GATGACTGGT TCAACAGCCA 700
200  TGGAACTAGG TGTGCAGGAG AAGTTTCTGC TGCAGCCAGC AACAATATCT 750
202  GTGGAGTCGG CGTAGCATAC AACTCCAAGG TGGCAGGGAT CCGGATGCTG 800
204  GACCAAGCCCT TTATGACAGA CATCATCGAA GCCTCCTCCA TCAGCCACAT 850
206  GCCTCAACTG ACAGTCACTT ACAGTGCAAG CTGGGGCCCC ACAGACAATG 900
208  GGAAGACGGT TGATGGGCCC CGAGAGCTCA CACTCCAGGC CATGGCTGAT 950
210  GGCCTGAACA AGGGCCGTGG GGGCAAAGGC AGCATCTATG TGTGGGCCTC 1000
212  TGGGGACGGT GGCAGCTACG ATGACTGCAA CTGTGACGGC TATGCTTCAA 1050
214  GCATGTGGAC CATCTCCATC AACTCAGCCA TCAATGATGG CAGGACTGCC 1100
216  TTGTATGATG AGAGTTGCTC TTCCACCTTA GCATCCACCT TCAGCAATGG 1150
218  GAGGAAGAGG AATCCTGAGG CTGGTGTGGC TACCACAGAC TTGTATGGCA 1200
220  ACTGTACTCT GAGACACTCT GGGACATCTG CAGCTGCTCC GGAGGCAGCT 1250
222  GCGGTGTTTG CATTAGCTTT GGAGGCTAAC CTGGATCTGA CCTGGAGAGA 1300
224  CATGCAACAT CTGACTGTGC TCACCTCCAA GCGGAACCAG CTTCATGATG 1350
226  AGGTTTCATCA GTGGCGACGG AATGGGGTTG GCCTGGAATT TAATCACCTC 1400
228  TTTGGCTACG GAGTCCTTGA TGCAGGTGCC ATGGTGAAAA TGGCTAAAGA 1450
230  CTGGAAGAACT GTTCCGGAGA GATTCCATTG TGTGGGAGGC TCTGTGCAGA 1500
232  ACCCTGAAAA AATACCACCC ACCGGCAAGC TGGTACTGAC CCTCAAAACA 1550
234  AATGCATGTG AGGGGAAGGA AAATTTCGTC CGTACCTGG AGCACGTCCA 1600
236  AGCTGTCAATC ACAGTCAACG CGACCAGGAG AGGAGACCTG AACATCAACA 1650
238  TGACCTCCCC AATGGGCACC AAGTCCATTT TGCTGAGCCG GCGTCCCAGA 1700
240  GACGACGACT CCAAGGTGGG CTTTGACAAG TGGCCTTTCA TGACCACCCA 1750
242  CACCTGGGGG GAGGATGCCC GAGGGACCTG GACCCTGGAG CTGGGGTTTG 1800
244  TGGCAGTGCC ACCACAGAAG GGGTTGCTGA AGGAATGGAC CCTGATGCTT 1850
246  CACGGCACAC AGAGCGCCCC ATACATCGAT CAGGTGGTGA GGGATTACCA 1900
248  GTCTAAGCTG GCCATGTCCA AGAAGCAGGA GCTGGAGGAA GAGCTGGATG 1950
250  AGGCTGTGGA GAGAAGCCTG CAAAGTATCC TGAGAAAGAA CTAGGGCCAC 2000
252  GCTTCCGAAT TC 2012

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/997,868

DATE: 03/25/2002

TIME: 15:23:32

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

254 (2) INFORMATION FOR SEQ ID NO: 3:

256 (i) SEQUENCE CHARACTERISTICS:

257 (A) LENGTH: 753 amino acids

258 (B) TYPE: Amino Acid

259 (D) TOPOLOGY: Linear

261 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

263	Met	Glu	Gln	Arg	Gly	Trp	Thr	Leu	Gln	Cys	Thr	Ala	Phe	Ala	Phe	
264	1				5				10							15
266	Phe	Cys	Val	Trp	Cys	Ala	Leu	Ser	Ser	Val	Lys	Ala	Lys	Arg	Gln	
267					20				25							30
269	Phe	Val	Asn	Glu	Trp	Ala	Ala	Glu	Ile	Pro	Gly	Gly	Gln	Glu	Ala	
270					35				40							45
272	Ala	Ser	Ala	Ile	Ala	Glu	Glu	Leu	Gly	Tyr	Asp	Leu	Leu	Gly	Gln	
273					50				55							60
275	Ile	Gly	Ser	Leu	Glu	Asn	His	Tyr	Leu	Phe	Lys	His	Lys	Ser	His	
276					65				70							75
278	Pro	Arg	Arg	Ser	Arg	Arg	Ser	Ala	Leu	His	Ile	Thr	Lys	Arg	Leu	
279					80				85							90
281	Ser	Asp	Asp	Asp	Arg	Val	Thr	Trp	Ala	Glu	Gln	Gln	Tyr	Glu	Lys	
282					95				100							105
284	Glu	Arg	Ser	Lys	Arg	Ser	Val	Gln	Lys	Asp	Ser	Ala	Leu	Asp	Leu	
285					110				115							120
287	Phe	Asn	Asp	Pro	Met	Trp	Asn	Gln	Gln	Trp	Tyr	Leu	Gln	Asp	Thr	
288					125				130							135
290	Arg	Met	Thr	Ala	Ala	Leu	Pro	Lys	Leu	Asp	Leu	His	Val	Ile	Pro	
291					140				145							150
293	Val	Trp	Glu	Lys	Gly	Ile	Thr	Gly	Lys	Gly	Val	Val	Ile	Thr	Val	
294					155				160							165
296	Leu	Asp	Asp	Gly	Leu	Glu	Trp	Asn	His	Thr	Asp	Ile	Tyr	Ala	Asn	
297					170				175							180
299	Tyr	Asp	Pro	Glu	Ala	Ser	Tyr	Asp	Phe	Asn	Asp	Asn	Asp	His	Asp	
300					185				190							195
302	Pro	Phe	Pro	Arg	Tyr	Asp	Leu	Thr	Asn	Glu	Asn	Lys	His	Gly	Thr	
303					200				205							210
305	Arg	Cys	Ala	Gly	Glu	Ile	Ala	Met	Gln	Ala	Asn	Asn	His	Lys	Cys	
306					215				220							225
308	Gly	Val	Gly	Val	Ala	Tyr	Asn	Ser	Lys	Val	Gly	Gly	Ile	Arg	Met	
309					230				235							240
311	Leu	Asp	Gly	Ile	Val	Thr	Asp	Ala	Ile	Glu	Ala	Ser	Ser	Ile	Gly	
312					245				250							255
314	Phe	Asn	Pro	Gly	His	Val	Asp	Ile	Tyr	Ser	Ala	Ser	Trp	Gly	Pro	
315					260				265							270
317	Asn	Asp	Asp	Gly	Lys	Thr	Val	Glu	Gly	Pro	Gly	Arg	Leu	Ala	Gln	
318					275				280							285
320	Lys	Ala	Phe	Glu	Tyr	Gly	Val	Lys	Gln	Gly	Arg	Gln	Gly	Lys	Gly	
321					290				295							300
323	Ser	Ile	Phe	Val	Trp	Ala	Ser	Gly	Asn	Gly	Gly	Arg	Gln	Gly	Asp	
324					305				310							315
326	Asn	Cys	Asp	Cys	Asp	Gly	Tyr	Thr	Asp	Ser	Ile	Tyr	Thr	Ile	Ser	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/997,868

DATE: 03/25/2002

TIME: 15:23:32

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

327		320		325		330
329	Ile Ser Ser Ala	Ser Gln Gln Gly Leu	Ser Pro Trp Tyr Ala	Glu		
330		335		340		345
332	Lys Cys Ser Ser	Thr Leu Ala Thr Ser	Tyr Ser Ser Gly Asp	Tyr		
333		350		355		360
335	Thr Asp Gln Arg	Ile Thr Ser Ala Asp	Leu His Asn Asp Cys	Thr		
336		365		370		375
338	Glu Thr His Thr	Gly Thr Ser Ala Ser	Ala Pro Leu Ala Ala	Gly		
339		380		385		390
341	Ile Phe Ala Leu	Ala Leu Glu Ala Asn	Pro Asn Leu Thr Trp	Arg		
342		395		400		405
344	Asp Met Gln His	Leu Val Val Trp Thr	Ser Glu Tyr Asp Pro	Leu		
345		410		415		420
347	Ala Ser Asn Pro	Gly Trp Lys Lys Asn	Gly Ala Gly Leu Met	Val		
348		425		430		435
350	Asn Ser Arg Phe	Gly Phe Gly Leu Leu	Asn Ala Lys Ala Leu	Val		
351		440		445		450
353	Asp Leu Ala Asp	Pro Arg Thr Trp Arg	Asn Val Pro Glu Lys	Lys		
354		455		460		465
356	Glu Cys Val Val	Lys Asp Asn Asn Phe	Glu Pro Arg Ala Leu	Lys		
357		470		475		480
359	Ala Asn Gly Glu	Val Ile Val Glu Ile	Pro Thr Arg Ala Cys	Glu		
360		485		490		495
362	Gly Gln Glu Asn	Ala Ile Lys Ser Leu	Glu His Val Gln Phe	Glu		
363		500		505		510
365	Ala Thr Ile Glu	Tyr Ser Arg Arg Gly	Asp Leu His Val Thr	Leu		
366		515		520		525
368	Thr Ser Ala Val	Gly Thr Ser Thr Val	Leu Leu Ala Glu Arg	Glu		
369		530		535		540
371	Arg Asp Thr Ser	Pro Asn Gly Phe Lys	Asn Trp Asp Phe Met	Ser		
372		545		550		555
374	Val His Thr Trp	Gly Glu Asn Pro Val	Gly Thr Trp Thr Leu	Lys		
375		560		565		570
377	Ile Thr Asp Met	Ser Gly Arg Met Gln	Asn Glu Gly Arg Ile	Val		
378		575		580		585
380	Asn Trp Lys Leu	Ile Leu His Gly Thr	Ser Ser Gln Pro Glu	His		
381		590		595		600
383	Met Lys Gln Pro	Arg Val Tyr Thr Ser	Tyr Asn Thr Val Gln	Asn		
384		605		610		615
386	Asp Arg Arg Gly	Val Glu Lys Met Val	Asn Val Val Glu Lys	Arg		
387		620		625		630
389	Pro Thr Gln Lys	Ser Leu Asn Gly Asn	Leu Leu Val Pro Lys	Asn		
390		635		640		645
392	Ser Ser Ser Ser	Asn Val Glu Gly Arg	Arg Asp Glu Gln Val	Gln		
393		650		655		660
395	Gly Thr Pro Ser	Lys Ala Met Leu Arg	Leu Leu Gln Ser Ala	Phe		
396		665		670		675
398	Ser Lys Asn Ala	Leu Ser Lys Gln Ser	Pro Lys Lys Ser Pro	Ser		
399		680		685		690

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/997,868

DATE: 03/25/2002

TIME: 15:23:33

Input Set : A:\P0748P3.txt

Output Set: N:\CRF3\03252002\I997868.raw

L:3 M:244 W: Invalid beginning of sequence listing, Data=[SEQUENCE LISTING], Duplicate Sequence Listing Title!

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:1112 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47